

# Profile Generation Methods for Reinforcing Robustness of Keystroke Authentication in Free Text Typing

Yoshitomo Matsubara<sup>1</sup>, Toshiharu Samura<sup>2</sup>, Haruhiko Nishimura<sup>1</sup>

<sup>1</sup>Graduate School of Applied Informatics, University of Hyogo, Kobe, Japan

<sup>2</sup>Department of Electrical and Computer Engineering, National Institute of Technology, Akashi College, Akashi, Japan

Email: [me1213@ob.akashi.ac.jp](mailto:me1213@ob.akashi.ac.jp), [samura@akashi.ac.jp](mailto:samura@akashi.ac.jp), [haru@ai.u-hyogo.ac.jp](mailto:haru@ai.u-hyogo.ac.jp)

Received 20 March 2015; accepted 13 April 2015; published 16 April 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

We have investigated several characteristics of the keystroke authentication in Japanese free text typing, and our methods have provided high recognition accuracy for high typing skill users who can type 700 or more letters per 5 minutes. There are, however, some situations decreasing the accuracy such as long period passage after registering each user's profile documents and existence of lower typing skill users who can type only about 500 - 600 letters per 5 minutes. In this paper, we propose new profile generation methods, profile-updating and profile-combining methods, to reinforce the robustness of keystroke authentication and show the effectiveness of them through three examinations with experimental data.

## Keywords

Keystroke Authentication, Biometrics, Information Security, Profile Generation, Robustness

---

## 1. Introduction

Keystroke data contain each user's fixed typing pattern, and a biometrics technology using such data is called keystroke authentication [1]-[3]. Comparing with other biometrics technology, keystroke authentication has two different features. First, keystroke data can be measured using only a keyboard; any special equipment, such as fingerprint and retinal scanners, is not required. Second, this biometric measuring system has possible applications other than access authorization, such as continuous observations of users on a computer system even after log-in process.

Most previous research on keystroke dynamics have focused on user authentication at the time of log-in, using not only information about a series of input characters for password recognition, but also keystroke authentication as a part of the authentication process [4] [5]. On the other hand, we considered the use of an analytic method that captures individual characteristics through the input of completely different phrases, rather than using repeated input of a short word for password verification. By using sentences of a certain length, it is possible to obtain sufficient information for deriving keystroke dynamics statistically. Little research has been performed on the keystroke authentication with such long-text input, and this has recently become the subject of academic discussion [6]-[13]. It can be used to detect spoofing after log-in process by observing keystroke data when using a system.

Using the WED + AD identification method which integrates the Weighted Euclidean Distance (WED) method proposed by Samura and Nishimura [10]-[13] and the Array Disorder (AD) method proposed by Gunetti and Picardi [7] [8], we could attain about 99% recognition accuracy for 127 high typing skill examinees who can type 700 or more letters per 5 minutes [11]. However, there remain some factors decreasing the recognition accuracy. In particular, long period passage after registration and existence of low typing skill users were found as significant factors through several experiments in our previous research [10]-[14]. Passage of time after registering each user's profile documents decreases the recognition accuracy, and lower typing skill examinees who can type only about 500 - 600 letters per 5 minutes can also deteriorate it.

Most approaches in the previous research on keystroke authentication have focused on how to treat feature extraction and identification methods, but hardly addressed issues on profile generation methods. In this paper, we propose new profile generation methods, profile-updating and profile-combining methods, in order to overcome the above mentioned degradation factors of the recognition accuracy. With the introduction of these methods, the effectiveness is demonstrated through three examinations with experimental data, by considering the findings that recognition accuracy is influenced by the number of registered profile documents, the number of kinds of letters and intervals between registration of profile documents, obtained in our works [14]-[16]. Implementation of the profile generation methods is expected to improve the recognition accuracy and reinforce the robustness of keystroke authentication system.

In Section 2, we briefly introduce the keystroke data and the feature extraction we have dealt with so far. Next, we present three methods for keystroke identification in free text typing in Section 3. In Section 4, new profile generation methods are proposed and examined with subjects' experimental data. Results and conclusion are given in Sections 5 and 6 respectively.

## 2. Keystroke Data and Feature Extraction

This section describes keystroke data collection system and how to extract features from keystroke data. In this study, we use a web-based system, and it has typing support software that is familiar to the participants, thereby lowering effects related to unfamiliarity and nervousness. **Figure 1** shows a screenshot of the software interface used in this study. The document display screen allows participants who are skilled typists to input text while viewing the Japanese text displayed in the upper row. Less skilled typists can type while viewing the Latin alphabet text displayed in the middle row. Latin alphabet text is removed from the screen as it is typed, allowing confirmation of mistyped characters. The top row displays the number of keystrokes, the number of errors and the amount of time remaining. Since this experiment focuses only on Latin alphabet input keystroke, conversion into Japanese kanji characters is not designed. Participants input a different text each time for 5 minutes. For each entry, typed key, key pressed time and key released time are recorded as raw data. **Figure 2** shows an example of collected keystroke data. The first, second and third fields show typed key, whether the key was pressed (p) or released (r) and UNIX time of the event.

Next, we describe feature extraction from keystroke data. The notation  $1pr$  in **Figure 3** indicates the time from press to release of a single key and is referred to below as key press duration. The notation  $2rr$  indicates the time from the release of one key to the release of the following key when typing a consonant-vowel pair. The time from release of the first key to the time of pressing the following key ( $2rp$ ) and the time from pressing the first key to pressing the second key ( $2pp$ ) are also considered. Furthermore,  $2pr1$  indicates the time from pressing the first key to releasing the next key,  $2pr2$  indicates the key press duration when typing the second (vowel) key, and  $2pr3$  indicates the key press duration when typing the first (consonant) key.

The average and standard deviation of each of the seven measures described above are used as the feature

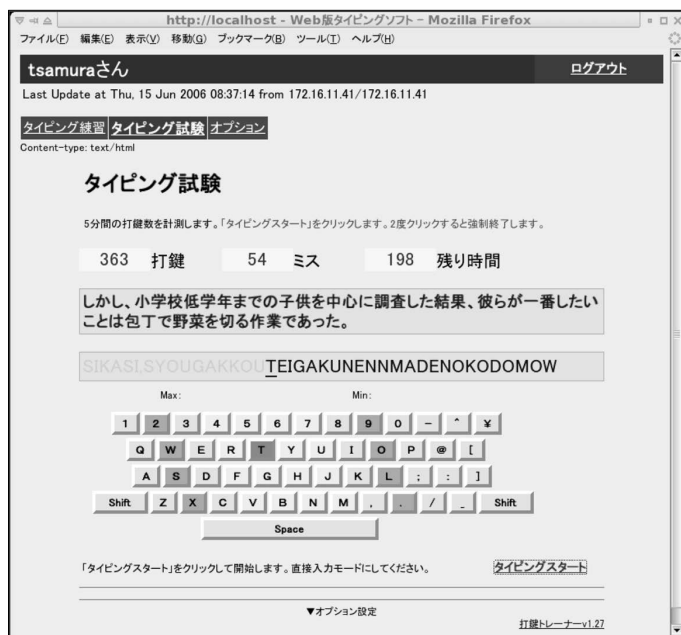


Figure 1. Snapshot of keystroke data collection system.

b	,	p	, 1197417770648
a	,	p	, 1197417770733
b	,	r	, 1197417770791
i	,	p	, 1197417770816
a	,	r	, 1197417770823
i	,	r	, 1197417770872
k	,	p	, 1197417770972
k	,	r	, 1197417771039
i	,	p	, 1197417771112
i	,	r	, 1197417771167

Figure 2. Example of collected keystroke data (typed key, pressed (p) or released (r), and UNIX time in order from the left).

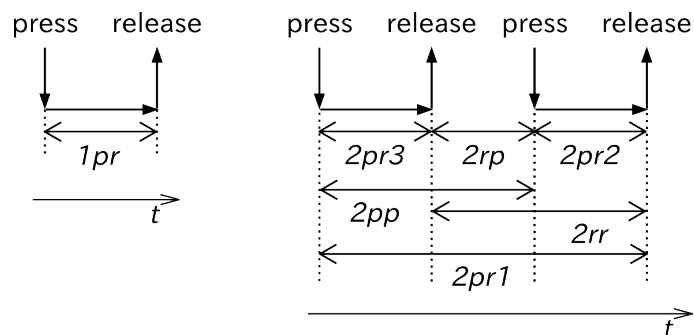


Figure 3. Keystroke measurements of single letter (left) and letter pair (right).

indices for identification of individuals. For these feature indices, standardization is performed according to the following equation:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (0 \leq x' \leq 1). \tag{1}$$

Here,  $x_{\min}$  and  $x_{\max}$  respectively refer to the minimum and maximum values obtained from the feature indices of all subjects. It has been confirmed in our previous works [10]-[13] that the standard deviations for six letter pairs (2\*\*) in feature indices have little contribution to improve recognition accuracy. Therefore, we do not adopt them in this study.

### 3. Identification Methods

This section describes identification methods: the Weighted Euclidean Distance (WED) method proposed by Samura and Nishimura [10]-[13], the Array Disorder (AD) method proposed by Gunetti and Picardi [7] [8] and WED + AD method proposed by Samura and Nishimura [13].

#### 3.1. Weighted Euclidean Distance (WED) Method

Taking the first profiling document of Typist A as  $docA1$ , the profile document of each participant can be represented as  $docA1, docA2, \dots, docAN, docB1, docB2, \dots, docBN, docC1, \dots$ . An unknown document is represented as  $docUK$ . The square root of WED used as the identification function is given by the following equation:

$$WED(docUK, docA1) = \sqrt{\frac{1}{m} \sum_{\alpha=1}^m \frac{1}{n_{\alpha}} \sum_{i=1}^{n_{\alpha}} (k_{\alpha(i)} - r_{\alpha(i)})^2} \quad (2)$$

The index of the feature indices  $\alpha (=1, 2, \dots)$  is then  $1pr.av, 1pr.sd, \dots$ . Furthermore,  $m$  is the number of contributing feature indices,  $\alpha(i)$  indicates the feature index  $\alpha$  for the  $i$ -th character (single letter or letter pair), and  $n_{\alpha}$  is the number of characters therein.  $n_{\alpha}$  will vary greatly with respect to the number of kinds of letters compared when, for example, taking the keystroke feature indices for single letters and those for two-letter combinations.  $k_{\alpha(i)}$  is the  $\alpha(i)$  feature index standardized according to Equation (1) for a profile document (e.g.,  $docA1$ ), and  $r_{\alpha(i)}$  is that for  $\alpha(i)$  of an unknown document ( $docUK$ ). The WED is normalized to the range 0 - 1 using the weightings  $\frac{1}{m}$  and  $\frac{1}{n_{\alpha}}$ .

Identification is performed using the nearest-neighbor rule. In other words, comparisons are performed between the unknown document and each of the profile documents, and the typist of the profile document that gives the lowest value is taken to be the typist of the unknown document. For example, given five profile documents (A1-A5) for a given subjects (Typist A), we can expect that if Typist A also typed the unknown document then its value will be close to one of the five states of Typist A A1-A5. In the case where the typist of the candidate profile document is matched with the typist of the test document, the identification is classified as a success; otherwise, the identification is classified as a failure. Such validation is performed in turn on each document used in this study, and the recognition accuracy is calculated by (number of successful identifications)/(number of test documents)  $\times$  100%.

#### 3.2. Array Disorder (AD) Method

The AD method, which is called the R-measure in refs. [7] [8], but referred to as the Array Disorder in the present study, ranks characters according to their feature index values, and evaluates the disorder of the rankings. Standardized feature indices are sorted in increasing order, the difference in rankings of each is taken, and the totals of each are taken as the distance.

$$AD(docUK, docA1) = \frac{1}{m} \sum_{\alpha=1}^m \frac{1}{\omega(n_{\alpha})} \sum_{i=1}^{n_{\alpha}} |rk_{\alpha(i)} - rr_{\alpha(i)}| \quad (3)$$

$$\omega(n_{\alpha}) = \begin{cases} \frac{n_{\alpha}^2}{2} & (n_{\alpha} : \text{even}) \\ \frac{n_{\alpha}^2 - 1}{2} & (n_{\alpha} : \text{odd}) \end{cases} \quad (4)$$

When  $n_{\alpha}$  characters are used to compare a feature index  $\alpha$ , if  $n_{\alpha}$  is even then the distance is divided by

$n_\alpha^2/2$ ; if  $n_\alpha$  is odd then the distance is divided by  $(n_\alpha^2 - 1)/2$ . Finally, the value is normalized to the range 0 - 1 by dividing the value by the number of contributing feature indices  $m$ .

### 3.3. WED + AD Method

In contrast to the WED method, which evaluates the magnitude of differences in feature index values between documents, the AD method focuses on differences between documents in ranking patterns of the feature indices. The WED + AD method complementarity incorporates the features of the WED and AD methods. The distance of the WED + AD method is given by the following equation:

$$[WED + AD](docUK, docA1) = WED(docUK, docA1) + AD(docUK, docA1). \quad (5)$$

In this method, neither *WED* nor *AD* dominates because they are normalized to the range 0 - 1. Detailed explanations for the above methods are given in ref. [13].

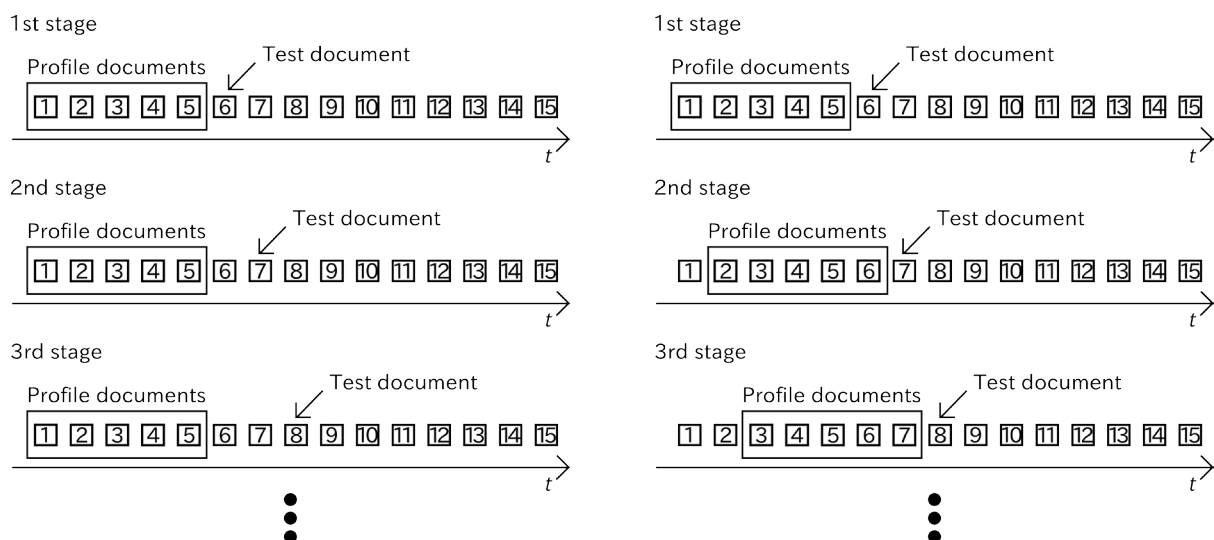
## 4. Profile Generation Methods and Examinations

### 4.1. Examination Using Profile-Updating Method

Since long period passage after registering profile documents decreases recognition accuracy [14], it is desired to prevent it from decreasing by removing an old profile document and adding a new profile document in order to operate a system sustainably. The profile-updating method removes the oldest profile document of all and registers a test document as a new profile document for the target user when the identification is succeeded. We show the effectiveness of this method through the following examination.

We choose 15 documents for each subject. Each document was collected almost every other week and sorted in time-series order. There are two cases of investigation: no-updating and updating profile documents. In the first case, five documents in order from the oldest are fixed as profile ones, and other 10 documents are used as test ones (no-update case). **Figure 4(left)** shows an example that there are 15 documents (profile documents: *doc1-doc5*, test documents: *doc6-doc15*). Identification is performed with sliding test document, *doc6*, *doc7*, ..., in order from the oldest. Other subjects' test documents are also used in the same way.

In the second case, we use the profile-updating method (update case). If the typist of the test document is the same with the typist of the nearest profile document with the test document, the test document is added as a new profile document to typist's profile documents, and the oldest one of them is removed. The test document is, however, not added and profile documents are not updated if the identification is failed. **Figure 4(right)** shows an example of using the profile-updating method. At first, *doc1-doc5* are profile documents, and *doc6* is the first test document and compared with profile documents. If identification of *doc6* is succeeded, *doc6* is registered as a new profile document and *doc1* is removed from profile documents. Second, *doc7* is com-



**Figure 4.** Identification processes in no-update (left) and update (right) cases.

pared with new profile documents. In this case, not only test documents but also profile documents are slidden if the identification is succeeded. Other subjects' documents are also used in the same way. In this examination, we take 100 subjects who have 15 or more documents, each of which contains 500 or more input letters. We evaluate recognition accuracy in both no-update and update cases.

### 4.2. Examination Using Profile-Combining Method

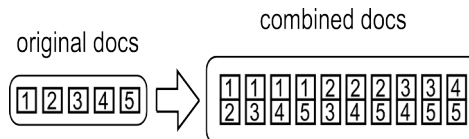
As for keystroke authentication in free text typing, high recognition accuracy is given if each document contain sufficient amounts of kinds of letters. In practice, it is desired to verify the user in a short time although it needs a long time to type such amount of letters. Therefore, we combine original profile documents (original docs) and generate new combined profile documents (combined docs). **Figure 5** shows an example of combining two profile documents for 5 original docs and generating 10 combined docs. Combining profile documents increases the number of input letters in each combined one. The more input letters, the more number of kinds of letters increases, and the sufficient number of kinds of letters improves the recognition accuracy [15].

In this examination, we prepare 5 or 10 profile documents (original docs) at random for each subject, and new profile documents (combined docs) are generated by combining two documents of them. **Figure 5** and **Figure 6** show the generation of new profile documents from 5 and 10 profile documents respectively. 10 ( $=_5C_2$ ) combined docs are made by 5 original docs for each subject as in **Figure 5**, and 45 ( $=_{10}C_2$ ) combined docs are made by 10 original docs as in **Figure 6**. Recognition accuracy is calculated by using the leave-one-out cross validation method. We choose a document as a test document from all documents, and calculate distances between the test document and the combined ones with each identification method. This is performed for all of documents. We do not compare a test document with a combined profile document if the combined profile document contains the test document. We take 250 subjects who have 10 or more documents, each of which contains 500 or more input letters. Among them, 68 subjects are the lower typing skill users who can type only less than 600 letters.

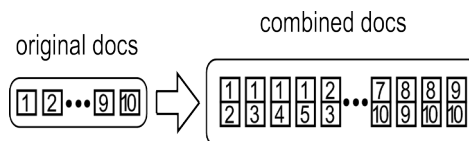
## 5. Results

As the WED + AD method gives the highest recognition accuracy among three identification methods [11] [13], we show only results with the WED + AD method in the following. **Figure 7** shows the recognition accuracy in both no-update and update cases explained in Subsection 4.1. Updating profile documents prevents the recognition accuracy from decreasing. Long period passage after registering profile documents decreases the recognition accuracy [14], and this is caused by the change of subjects' typing patterns change over time. In fact, **Figure 8** shows the increment of the average number of input letters for 100 subjects in each-week documents, and it indicates that their typing patterns changed as time passed. Updating profile documents keeps the period after registration short and enables the system to adapt the changes of typing patterns in their test documents

**Figure 9** shows the recognition accuracy for combined profile documents constructed by the method in Subsection 4.2. The recognition accuracy in 10-document case is about 4% - 5% higher than that in 5-document case. This is owing to statistical precision enhancement by the rise in the number of documents. In addition,



**Figure 5.** Generation of new combined profile documents from 5 original profile documents.



**Figure 6.** Generation of new combined profile documents from 10 original profile documents.

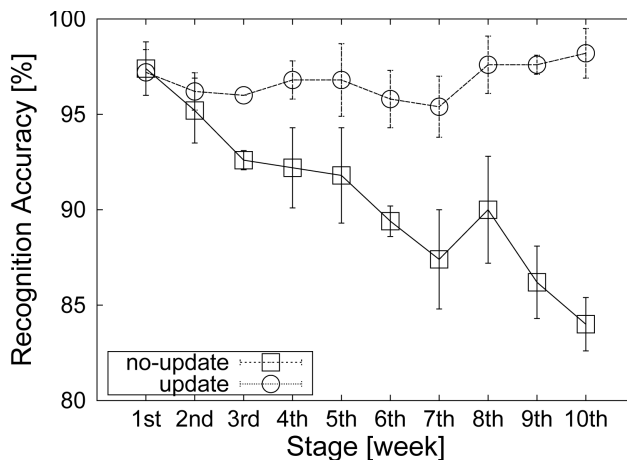


Figure 7. Recognition accuracy in no-update and update cases.

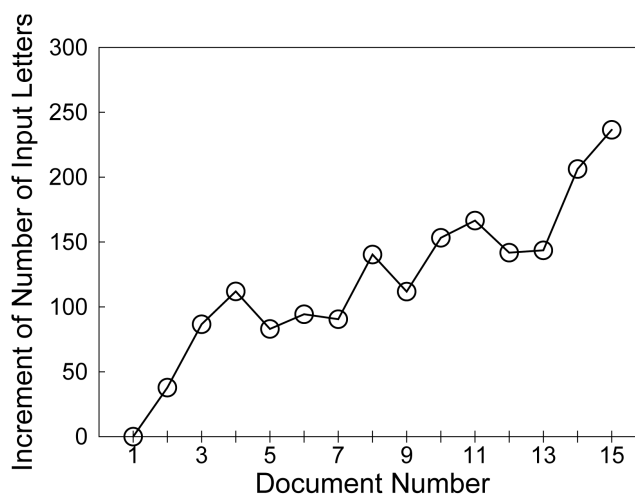


Figure 8. Increment of average number of input letters for 100 subjects in each-week document.

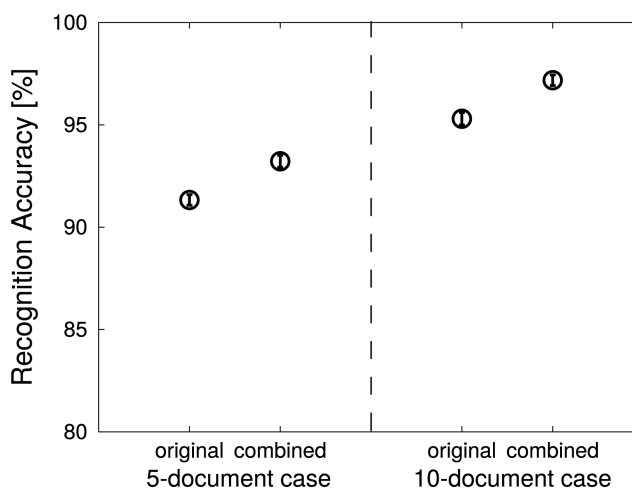
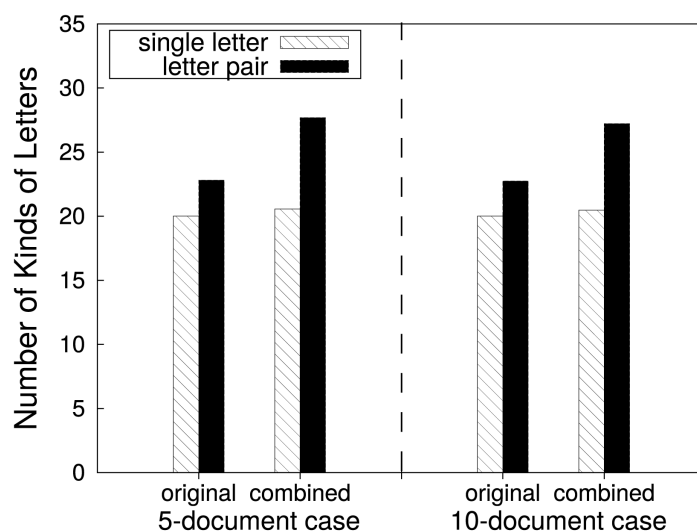


Figure 9. Comparison of recognition accuracy between original profile documents and combined ones (left: 5-document case, right: 10-document case).

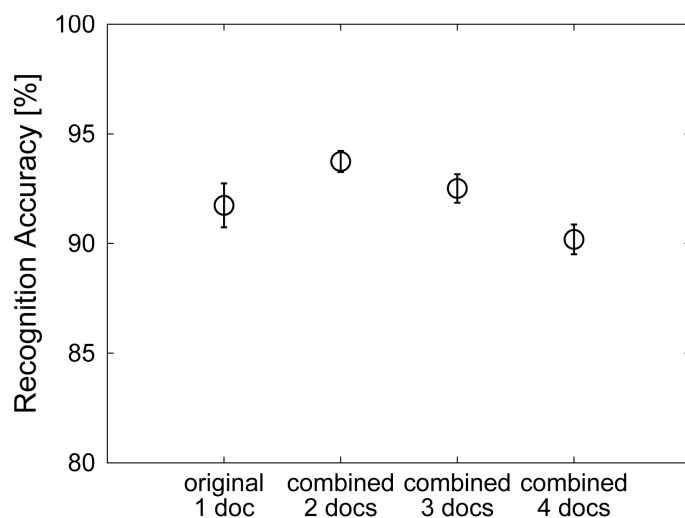
comparing with the results of original documents, combining profile documents is found to improve the recognition accuracy even for the case of including lower typing skill subjects who can type only about 500 - 600 letters per 5 minutes (68 subjects). And then, both in the 5-document and 10-document cases, their values of recognition accuracy increase a few percent comparably by combining profile documents.

**Figure 10** shows the comparison of the number of kinds of letters,  $n_a$ , in Equations (2), (3) and (4) per one profile document corresponding to each in **Figure 9**. There is little difference in the number of kinds of single letters between the original and combined profile documents because it is rich even in the original profile documents, but the number of kinds of letter pairs increases by combining profile documents. This means a dimensional extension in the contributing feature indices which leads to a better recognition accuracy.

In the case in **Figure 9**, two profile documents are used to generate one combined profile document, but it is possible to use more than two ones as constituent documents for that. **Figure 11** shows the recognition accuracy obtained in each case using different number of constituent profile documents for a combined profile document in 250 subjects with 5 original profile documents. When the number of constituent profile documents is two, it provides the highest recognition accuracy.



**Figure 10.** Change of the number of kinds of letters per one profile document.



**Figure 11.** Dependence of recognition accuracy on the number of constituent profile documents for combining.



Figure 12 shows the change of the number of kinds of letters  $n_a$  corresponding to each in Figure 11. The number of kinds of single letters attains its maximum even at the stage of one original profile documents, and the number of kinds of letter pairs increases and reaches near the plateau at the stage of four constituent profile documents. On the other hand, the number of newly produced combined profile documents from 5 original profile documents becomes 6 ( $= {}_{5-1}C_2$ ) for two constituents, 4 ( $= {}_{5-1}C_3$ ) for three constituents, and 1 ( $= {}_{5-1}C_4$ ) for four constituents. Note that when we use the leave-one-out cross validation for a test document to be checked, combined profile documents which contain the exact test document should be removed from the group of them. Based on the above results, both the numbers of kinds of letters and valid combined profile documents increase only in the case of combining two constituent profile documents and this enhances the recognition accuracy as shown in Figure 11. In other cases, the recognition accuracy has a deteriorative tendency due to the trade-off relationship between the two factors.

Through the examinations in this paper, it has been demonstrated that the profile-updating method prevents the recognition accuracy from decreasing, and the profile-combining method improves it. By adopting both methods together as shown in Figure 13, we can expect the system to sustain higher performance. Figure 14 shows the recognition accuracy for the same 100 subjects in examination in Subsection 4.1. By applying the profile updating and combining methods simultaneously (update + combine), the recognition accuracy is found to keep the higher value around 98% in comparison with the result using only the profile-updating method.

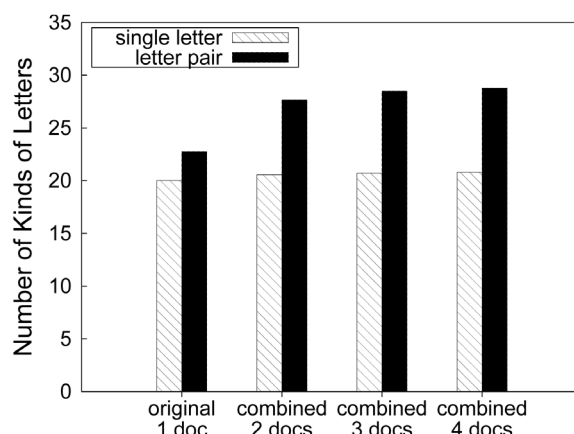


Figure 12. Dependence of the number of kinds of letters on the number of constituent profile documents for combining.

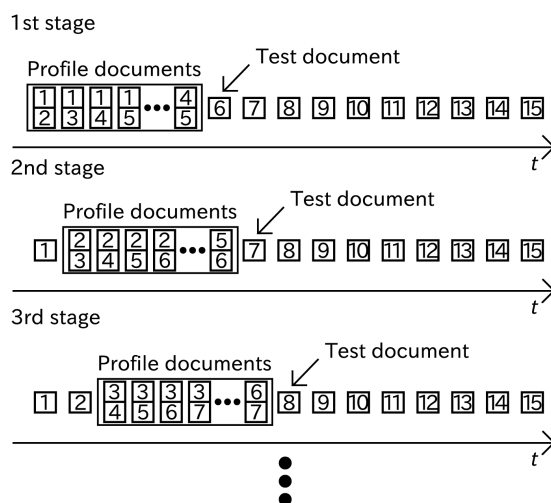
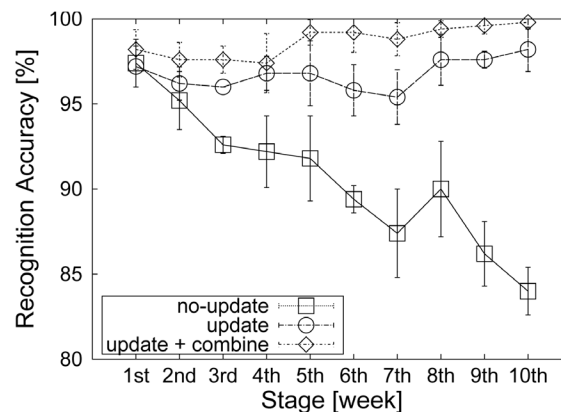


Figure 13. Identification process adopting both profile updating and combining methods for 5 original profile documents (update + combine case).



**Figure 14.** Recognition accuracy in update + combine case compared with the result in no-update and update cases in Figure 7.

## 6. Conclusions

In this paper, we have proposed two new profile generation methods to reinforce the robustness of keystroke authentication system for free text typing and shown the effectiveness of them through the examinations with subjects' practical data. The profile-updating method kept the period after registration short and enabled the system to adapt the changes of users' typing patterns, and thus prevented the recognition accuracy from decreasing. In addition, the profile-combining method increased both the numbers of kinds of letters and combined profile documents and owing to their synergistic effect, improved the recognition accuracy even in the case of including lower typing skill subjects. Furthermore, adopting both methods together sustained the high recognition accuracy and reinforced the robustness of keystroke authentication in free text typing.

In general, these profile generation methods are independent of feature extraction scheme in Section 2 and identification methods in Section 3 and thus can be widely applied to other keystroke authentication approaches. Although we have investigated on Japanese text input in this study, the feature indices and authentication model used here can be applied to other language text input.

## Acknowledgments

This work was supported by the Grant in Aid for Scientific Research (C24500101) from JSPS.

## References

- [1] Banerjee, S. and Woodard, D. (2012) Biometric Authentication and Identification Using Keystroke Dynamics: A Survey. *Journal of Pattern Recognition Research*, **7**, 116-139. <http://dx.doi.org/10.13176/11.427>
- [2] Alsultan, A. and Warwick, K. (2013) Keystroke Dynamics Authentication: A Survey of Free-text Methods. *International Journal of Computer Science Issues*, **10**, 1-10.
- [3] Teh, P.S., Teoh, A.B.J. and Yue, S. (2013) A Survey of Keystroke Dynamics Biometrics. *The Scientific World Journal*, **2013**, Article ID: 408280, 24 p.
- [4] Umphress, D. and Williams, G. (1985) Identity Verification through Keyboard Characteristics. *International Journal of Man-Machine Studies*, **23**, 263-273. [http://dx.doi.org/10.1016/S0020-7373\(85\)80036-5](http://dx.doi.org/10.1016/S0020-7373(85)80036-5)
- [5] Joyce, R. and Gupta, G. (1990) Identity Authentication Based on Keystroke Latencies. *Communications of the ACM*, **33**, 168-176. <http://dx.doi.org/10.1145/75577.75582>
- [6] Monroe, F. and Rubin, A. (2000) Keystroke Dynamics as a Biometric for Authentication. *Future Generation Computer Systems*, **16**, 351-359. [http://dx.doi.org/10.1016/S0167-739X\(99\)00059-X](http://dx.doi.org/10.1016/S0167-739X(99)00059-X)
- [7] Bergadano, F., Gunetti, D. and Picardi, C. (2002) User Authentication through Keystroke Dynamics. *ACM Transactions of Information and Systems Security*, **5**, 367-397. <http://dx.doi.org/10.1145/581271.581272>
- [8] Gunetti, D. and Picardi, C. (2005) Keystroke Analysis of Free Text. *ACM Transactions of Information and System Security*, **8**, 312-347. <http://dx.doi.org/10.1145/1085126.1085129>
- [9] Tappert, C.C., Villani, M. and Cha, S. (2009) Keystroke Biometric Identification and Authentication on Long-Text In-

- 
- put. *Behavioral Biometrics for Human Identification: Intelligent Applications*, IGI Global, 342-367.
- [10] Samura, T. and Nishimura, H. (2009) Personal Identification by Keystroke Dynamics in Japanese Free Text Typing. *Transactions of the Institute of Systems, Control and Information Engineers*, **22**, 145-153. [In Japanese]  
<http://dx.doi.org/10.5687/iscie.22.145>
- [11] Samura, T. and Nishimura, H. (2010) A Hybrid Model for Individual Identification Based on Keystroke Data in Japanese Free Text Typing. *Transactions of the Society of Instrument and Control Engineers*, **46**, 676-684. [In Japanese]  
<http://dx.doi.org/10.9746/sicetr.46.676>
- [12] Samura, T. and Nishimura, H. (2011) Keystroke Dynamics for Individual Identification in Japanese Free Text Typing. *SICE Journal of Control, Measurement, and System Integration*, **4**, 172-176.  
<http://dx.doi.org/10.9746/jcmsi.4.172>
- [13] Samura, T. and Nishimura, H. (2011) Personal Identification and Authentication Based on Keystroke Dynamics in Japanese Long-Text. In: Traore, I., et al., Eds., *Continuous Authentication based on Biometrics, Data, Models, and Metrics*, IGI Global, 212-231.
- [14] Matsubara, Y., Samura, T. and Nishimura, H. (2012) Robustness of Keystroke Dynamics against Changing in Input and Recognition Conditions for Japanese Atypical Text. *Proceedings of 2012 International Symposium on Technology for Sustainability*, Bangkok, November 2012, 103-106.
- [15] Hiraoka, H., Ishii, M., Samura, T. and Nishimura, H. (2011) Keystroke Dynamics in Japanese Free text Typing. *Proceedings of Forum on Information Technology*, Hakodate, September 2011, 19-24. [In Japanese]
- [16] Samura, T., Matsubara, Y. and Nishimura, H. (2013) Performance Assessment in Keystroke Dynamics by Combined Profile Documents for Free Text Typing. *Proceedings of SICE Annual Conference*, 2013, 265-270.